UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a.	3. DISTRIBUTION / AVAILABILITY OF REPORT
2b.	Approved for public release; distribution unlimited.
4. AD-A205 040	5. MONITORING ORGANIZATION REPORT NUMBER(S)
	AFOSR-TR- 89-0191
6a. NAME OF PERFORMING ORGANIZATION: University of Pennsylvania, Computer & Inform. Science Dept.	6b. OFFICE SYMBOL (If applicable):
	7a. NAME OF MONITORING ORGANIZATION: AFOSR
6c. ADDRESS (City, State, and ZIP Code): Philadephhia, PA	7b. ADDRESS (City, State, and ZIP Code): Building 410, Bolling, AFB DC 20332-6448
8a. NAME OF FUNDING / SPONSORING ORGANIZATION: AFOSR	8b. OFFICE SYMBOL (If applicable): NM
	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR 87-0186
8c. ADDRESS (City, State, and ZIP Code): Building 410, Bolling, AFB DC 20332-6448	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61102F	2304	A2	

11. TITLE (Include Security Classification)

LANDSCAN: Query Driven Recognition System

12. PERSONAL AUTHOR(S)
Dr. Ruzena Bajcsy

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM Apr 1987 TO Mar'88	March 1988	3

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

LandScan was intended to be an interface not to a database but to an active (and interactive) visual recognition system. That is, rather than searching a body of exiting facts about the donain, the system drives a vision component that will process data supplied to it by two cameras and respond with identification and analysis of objects found in the scene. We currently are looking at a scale model of a city block that is part of the University of Pennsylvania campus. Obviously, knowledge about languaga, the world and visual properties of objects is needed for this, and will have to reside in the various components of the system, but the system will be gathered in response to the user's requests.

DTIC ELECTE 23 FEB 1989 S D

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Dr. Abraham Waksman	(202) 767-5207	NM

DD Form 1473, JUN 86 — Previous editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Final Report
Landscan: Query Driven Recognition System

April 1987 - March 1988

Dr. Ruzena Bajcsy
Computer and Information Science Department
University of Pennsylvania
Philadelphia, PA 19104

1. Summary of Work

During the last four years, through AFOSR support, we have been privileged to have the opportunity to work on the Natural Language Driven Scene Analyzer known by the acronym ***Landscan***.

LandScan was intended to be an interface not to a database but to an active (and interactive) visual recognition system. That is, rather than searching a body of existing facts about the domain, the system drives a vision component that will process data supplied to it by two cameras and respond with identification and analysis of objects found in the scene. We currently are looking at a scale model of a city block that is part of the University of Pennsylvania campus. Obviously, knowledge about language, the world, and visual properties of objects is needed for this, and will have to reside in the various components of the system, but the data being returned by the system will be gathered in response to the user's requests.

One of the primary motivations of this research was to examine how the questions posed by the user and transmitted to the vision system by the interface can constrain/limit the amount of image processing done. What we have found out, however, was that the questions that were interesting to the natural language people, like Herskovits, implied very sophisticated visual recognition that we, the vision people, could not deliver. In particular, the natural language researchers were interested in various details, such as discriminating between different types of roofs and buildings, whereas the vision researchers were facing problems of errors and mistakes that were due to poor illumination, non-robust edge detection, sparse data obtained by stereo algorithms, and in general, inconsistent segmentation of the scene that was to be interpreted by the reasoner coming from the natural language end.

This realisation redirected the focus of our research in the vision area for the past year. (which started in March 1987 and continues until March 1988) We have spent this time thoroughly examining the parameters of the segmentation process.

What we have learned is:

1. Segmentation can be casted as a global process (performed on the whole picture) with some given external parameters, i.e., topological parameters that are also the goal of this process, and the noise of the camera system. This global process invokes some local (neighborhood bound) processes such as edge detection and boundary formation and region growing, or grouping, of similar elements. These local processes have also parameters which can be built in a priori or can be variable and computed by the segmentation processor. These parameters are: size of the desired bandpass for filtering and the range of the contrast so that proper discontinuities (edges) can be detected and the local similarity in order to detect continuities (regions).

2. Secondly, we are learning how the local parameters interact with the global parameters for a given goal of the segmentation process. Clearly the result of the segmentation, i.e., the segmented image is not *unique*, and hence, must be described in addition to its geometry with some degree of uncertainty. This is well justified since the problem is unconstrained, there is just not enough information (bottom-up or topdown). In fact, this is an example of the famous psychological problem of Figure-Ground disambiguation.

2. Publications

Anderson, Helen L., "Edge Detection for Obejct Recognition in Aerial Photographs", Department of Computer and Information Science Technical Report MS-CIS-87-14.

Anderson, Helen L., Bajcsy, R., Mintz, M., "A Modular Feedback System for Image Segmentation", Department of Computer and Information Science Technical Report MS-CIS-8

Anderson, Helen L., Bajcsy, R., Mintz, M., "A Modular Feedback System for Image Segementation", Department of Computer and Information Science Technical Report MS-CIS-87-56.

Krotkov, Eric P., "Exploratory Visual Sensing for Determining Spatial Layout with an Agile Stereo Camera System", University of Pennsylvania Ph.D. Dissertation, 1987.

Krotkov, Eric P. and Kories, R., "Integrating Multiple Uncertain View of a Static Scene Acquired by an Agile Camera System", Department of Computer and Information Science Technical Report MS-CIS-88-11.

Krotkov, Eric P. and Kories, R., "Cooperative Focus and Stereo Ranging" Department of Computer and Information Science Technical Report MS-CIS-88-88.